



COPY OF PAPERS
ORIGINALLY FILED

SEQUENCE LISTING

#6

GENERAL INFORMATION

<110> APPLICANT: Lingappa, Jaisri
Lingappa, Vishwanath

<120> TITLE OF THE INVENTION: HIV Capsid Assembly Associated
Compositions and Method

<130> FILE REFERENCE: UCSF.002.01US

<140> CURRENT APPLICATION NUMBER: 10/040,206

<141> CURRENT FILING DATE: 2002-01-02

CORRESPONDENCE ADDRESS:

ADDRESSEE: Rae-Venter Law Group, P.C.

STREET: P.O. Box 1898

CITY: Monterey

STATE: CA

COUNTRY: US

ZIP: 93942-1898

ATTORNEY/AGENT INFORMATION:

NAME: Barbara Rae-Venter

REGISTRATION NUMBER: 32,750

TELECOMMUNICATION INFORMATION:

TELEPHONE: 831-648-3090

TELEFAX: 831-242-0137

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: WINDOWS

SOFTWARE: PatentIn version 3.1

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 09/020,144

FILING DATE: 06-FEB-1998

<160>NUMBER OF SEQUENCES: 6

<170> PatentIn Version 3.0

<210>INFORMATION FOR SEQ ID NO:1

SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1610 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

(C) ISOLATE: DNA coding sequence for HIV capsid protein Pr55

<400> SEQUENCE DESCRIPTION: SEQ ID NO:1

ATGGGTGCGA	GAGCGTCGGT	ATTAAGCGGG	GGAGAATTAG	ATAAATGGGA	AAAAATTCGG	60
TTAAGGCCAG	GGGGAAAGAA	AAAATATAAG	TTAAAACATA	TAGTATGGGC	AAGCAGGGAG	120
CTAGAACGAT	TCGCAGTCAA	TCCTGGCCTG	TTAGAAACAT	CAGAAGGCTG	CAGACAAATA	180
TTGGGACAGC	TACAGCCATC	CCTTCAGACA	GGATCAGAAG	AACTTAGATC	ATTATATAAT	240
ACAGTAGCAA	CCCTCTATTG	TGTACATCAA	AGGATAGATG	TAAAAGACAC	CAAGGAAGCT	300
TTAGAGAAGA	TAGAGGAAGA	GCAAAACAAA	AGTAAGAAAA	AGGCACAGCA	AGCAGCAGCT	360
GCAGCTGGCA	CAGGAAACAG	CAGCCAGGTC	AGCCAAAATT	ACCCTATAGT	GCAGAACCTA	420
CAGGGGCAAA	TGGTACATCA	GGCCATATCA	CCTAGAACTT	TAAATGCATG	GGTAAAAGTA	480
GTAAGAAGAA	AGGCTTTCAG	CCCAGAAGTA	ATACCCATGT	TTTCAGCATT	ATCAGAAGGA	540
GCCACCCAC	AAGATTTAAA	CACCATGCTA	AACACAGTGG	GGGGACATCA	AGCAGCCATG	600
CAAATGTTAA	AAGAGACTAT	CAATGAGGAA	GCTGCAGAAT	GGGATAGAGT	GCATCCAGTG	660
CATGCAGGGC	CTATTGCACC	AGGCCAAATG	AGAGAACCAA	GGGGAAGTGA	CATAGCAGGA	720
ACTACTAGTA	CCCTTCAGGA	ACAAATAGGA	TGGATGACAA	ATAATCCACC	TATCCCAGTA	780
GGAGAAATCT	ATAAAAGATG	GATAATCCTG	GGATTAAATA	AAATAGTAAG	AATGTATAGC	840
CCTACCAGCA	TTCTGGACAT	AAGACAAGGA	CCAAAGGAAC	CCTTTAGAGA	TTATGTAGAC	900
CGGTTCTATA	AAACTCTAAG	AGCCGAACAA	GCTTCACAGG	ATGTAAAAAA	TTGGATGACA	960
GAAACCTTGT	TGGTCCAAAA	TGCAAACCCA	GATTGTAAGA	CTATTTTAAA	AGCATTGGGA	1020
CCAGCAGCTA	CACTAGAAGA	AATGATGACA	GCATGTCAGG	GAGTGGGGGG	ACCCGGCCAT	1080
AAAGCAAGAG	TTTTGGCTGA	AGCCATGAGC	CAAGTAACAA	ATCCAGCTAA	CATAATGATG	1140
CAGAGAGGCA	ATTTTAGGAA	CCAAAGAAAG	ACTGTTAAGT	GTTTCAATTG	TGGCAAAGAA	1200

(C) ISOLATE: Degenerate oligonucleotide C-terminal peptide
sequence of WGHP68

<400> SEQUENCE DESCRIPTION: SEQ ID NO:3

ATGAATTCAC TGGGACTGCG GATAGATTAC TGGTACTGGG GATC 44

<210> INFORMATION FOR SEQ ID NO:4

SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 44 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: cDNA

(vi) ORIGINAL SOURCE:

(C) ISOLATE: Degenerate oligonucleotide C-terminal peptide
sequence of WGHP68

<400> SEQUENCE DESCRIPTION: SEQ ID NO:4

ATGAATTCAC TGGGCTCTGA TAGATTACTG GTACTGGGGA TC 42

<210> SEQ ID NO:5

<211> Length:604

<212> Type: PRT

<213> Organism: Triricum aestivum

<400> Sequence 5

Met	Ala	Asp	Arg	Leu	Thr	Arg	Ile	Ala	Ile	Val	Ser	Glu	Asp	Lys	Cys
1				5				10						15	
Lys	Pro	Lys	Lys	Cys	Arg	Gln	Glu	Cys	Lys	Lys	Ser	Cys	Pro	Val	Val
				20				25						30	
Lys	Thr	Gly	Lys	Leu	Cys	Ile	Glu	Val	Ser	Pro	Val	Ala	Lys	Leu	Ala
				35				40						45	
Phe	Ile	Ser	Glu	Glu	Leu	Cys	Ile	Gly	Cys	Gly	Ile	Cys	Val	Lys	Lys

50	55	60
Cys Pro Phe Asp Ala Ile Glu Ile Ile Asn Leu Pro Lys Asp Leu Glu		
65	70	75
Lys Asp Thr Thr His Arg Tyr Gly Pro Asn Thr Phe Lys Leu His Arg		
	85	90
Leu Pro Val Pro Arg Pro Gly Gln Val Leu Gly Leu Val Gly Thr Asn		
	100	105
Gly Ile Gly Lys Ser Thr Ala Leu Lys Val Leu Ala Gly Lys Leu Lys		
	115	120
Pro Asn Leu Gly Arg Phe Lys Asn Pro Pro Asp Trp Gln Glu Ile Leu		
	130	140
Thr Tyr Phe Arg Gly Ser Glu Leu Gln Asn Tyr Phe Thr Arg Ile Leu		
145	150	155
Glu Asp Asn Leu Lys Ala Ile Ile Lys Pro Gln Tyr Val Asp His Ile		
	165	170
Pro Lys Ala Val Gln Gly Asn Val Gly Gln Val Leu Glu Gln Lys Asp		
	180	185
Glu Arg Asp Met Lys Asn Glu Leu Cys Val Asp Leu Glu Leu Asn Gln		
	195	200
Val Ile Asp Arg Asn Val Gly Asp Leu Ser Gly Gly Glu Leu Gln Arg		
	210	215
Phe Ala Ile Ala Val Val Ala Val Gln Ser Ala Glu Ile Tyr Met Phe		
225	230	235
Asp Glu Pro Ser Ser Tyr Leu Asp Val Lys Gln Arg Leu Lys Ala Ala		
	245	250
Arg Val Ile Arg Ser Leu Leu Arg Ser Asn Ser Tyr Val Ile Val Val		
	260	265
Glu His Asp Leu Ser Val Leu Asp Tyr Leu Ser Asp Phe Ile Cys Cys		
	275	280
Leu Tyr Gly Lys Pro Gly Ala Tyr Gly Val Val Thr Leu Pro Phe Ser		
	290	295
Val Arg Glu Gly Ile Asn Ile Phe Leu Ala Gly Phe Val Pro Thr Glu		
305	310	315
Asn Leu Arg Phe Arg Asp Glu Ser Leu Thr Phe Lys Ile Ala Glu Thr		
	325	330
Gln Glu Ser Ala Glu Glu Val Ala Thr Tyr Gln Arg Tyr Lys Tyr Pro		
	340	345
Thr Met Ser Lys Thr Gln Gly Asn Phe Lys Leu Ser Val Val Glu Gly		
		350

355	360	365
Glu Phe Thr Asp Ser Gln Ile Val Val Met Leu Gly Glu Asn Gly Thr		
370	375	380
Gly Lys Thr Thr Phe Ile Arg Met Leu Ala Gly Leu Leu Lys Pro Asp		
385	390	395
Thr Met Glu Gly Thr Glu Val Glu Ile Pro Glu Phe Asn Val Ser Tyr		
405	410	415
Lys Pro Gln Lys Ile Ser Pro Lys Phe Gln His Pro Val Arg His Leu		
420	425	430
Leu His Ser Lys Ile Arg Asp Ser Tyr Thr His Pro Gln Phe Val Ser		
435	440	445
Asp Val Met Lys Pro Leu Gln Ile Glu Gln Leu Met Asp Gln Glu Val		
450	455	460
Ile Asn Leu Ser Gly Gly Glu Leu Gln Arg Val Ala Leu Cys Leu Cys		
465	470	475
Leu Gly Lys Pro Ala Asp Ile Tyr Leu Ile Asp Glu Pro Ser Ala Tyr		
485	490	495
Leu Asp Ser Glu Gln Arg Ile Val Ala Ser Lys Val Ile Lys Arg Phe		
500	505	510
Ile Leu His Ala Lys Lys Thr Ala Phe Ile Val Glu His Asp Phe Ile		
515	520	525
Met Ala Thr Tyr Leu Ala Asp Lys Val Ile Val Tyr Glu Gly Leu Ala		
530	535	540
Ser Ile Asp Cys Thr Ala Asn Ala Pro Gln Ser Leu Val Ser Gly Met		
545	550	555
Asn Lys Phe Leu Ser His Leu Asp Ile Thr Phe Arg Arg Asp Pro Thr		
565	570	575
Asn Tyr Arg Pro Arg Ile Asn Lys Leu Glu Ser Thr Lys Asp Arg Glu		
580	585	590
Gln Lys Asn Ala Gly Ser Tyr Tyr Tyr Leu Asp Asp		
595	600	

<210> SEQ ID NO:6

<211> LENGTH: 599

<212> TYPE: PRT

<213> ORGANISM: Homo sapiens

<400> SEQUENCE 6

Met	Ala	Asp	Lys	Leu	Thr	Arg	Ile	Ala	Ile	Val	Asn	His	Asp	Lys	Cys
1				5					10					15	
Lys	Pro	Lys	Lys	Cys	Arg	Gln	Glu	Cys	Lys	Lys	Ser	Cys	Pro	Val	Val
			20					25					30		
Arg	Met	Gly	Lys	Leu	Cys	Ile	Glu	Val	Thr	Pro	Gln	Ser	Lys	Ile	Ala
		35					40					45			
Trp	Ile	Ser	Glu	Thr	Leu	Cys	Ile	Gly	Cys	Gly	Ile	Cys	Ile	Lys	Lys
	50					55					60				
Cys	Pro	Phe	Gly	Ala	Leu	Ser	Ile	Val	Asn	Leu	Pro	Ser	Asn	Leu	Glu
65					70				75					80	
Lys	Glu	Thr	Thr	His	Arg	Tyr	Cys	Ala	Asn	Ala	Phe	Lys	Leu	His	Arg
				85					90					95	
Leu	Pro	Ile	Pro	Arg	Pro	Gly	Glu	Val	Leu	Gly	Leu	Val	Gly	Thr	Asn
			100					105					110		
Gly	Ile	Gly	Lys	Ser	Ala	Ala	Leu	Lys	Ile	Leu	Ala	Gly	Lys	Gln	Lys
		115					120					125			
Pro	Asn	Leu	Gly	Lys	Tyr	Asp	Asp	Pro	Pro	Asp	Trp	Gln	Glu	Ile	Leu
	130					135					140				
Thr	Tyr	Phe	Arg	Gly	Ser	Glu	Leu	Gln	Asn	Tyr	Phe	Thr	Lys	Ile	Leu
145					150					155				160	
Glu	Asp	Asp	Leu	Lys	Ala	Ile	Ile	Lys	Pro	Gln	Tyr	Val	Ala	Arg	Phe
			165						170					175	
Leu	Arg	Leu	Ala	Lys	Gly	Thr	Val	Gly	Ser	Ile	Leu	Asp	Arg	Lys	Asp
		180						185					190		
Glu	Thr	Lys	Thr	Gln	Ala	Ile	Val	Cys	Gln	Gln	Leu	Asp	Leu	Thr	His
		195					200						205		
Leu	Lys	Glu	Arg	Asn	Val	Glu	Asp	Leu	Ser	Gly	Gly	Glu	Leu	Gln	Arg
	210					215					220				
Phe	Ala	Cys	Ala	Val	Val	Cys	Ile	Gln	Lys	Ala	Asp	Ile	Phe	Met	Phe
225				230						235				240	
Asp	Glu	Pro	Ser	Ser	Tyr	Leu	Asp	Val	Lys	Gln	Arg	Leu	Lys	Ala	Ala
			245						250					255	
Ile	Thr	Ile	Arg	Ser	Leu	Ile	Asn	Pro	Asp	Arg	Tyr	Ile	Ile	Val	Val
		260						265					270		
Glu	His	Asp	Leu	Ser	Val	Leu	Asp	Tyr	Leu	Ser	Asp	Phe	Ile	Cys	Cys
		275						280					285		
Leu	Tyr	Gly	Val	Pro	Ser	Ala	Tyr	Gly	Val	Val	Thr	Met	Pro	Phe	Ser
	290					295							300		

Val Arg Glu Gly Ile Asn Ile Phe Leu Asp Gly Tyr Val Pro Thr Glu
 305 310 315 320
 Asn Leu Arg Phe Arg Asp Ala Ser Leu Val Phe Lys Val Ala Glu Thr
 325 330 335
 Ala Asn Glu Glu Glu Val Lys Lys Met Cys Met Tyr Lys Tyr Pro Gly
 340 345 350
 Met Lys Lys Lys Met Gly Glu Phe Glu Leu Ala Ile Val Ala Gly Glu
 355 360 365
 Phe Thr Asp Ser Glu Ile Met Val Met Leu Gly Glu Asn Gly Thr Gly
 370 375 380
 Lys Thr Thr Phe Ile Arg Met Leu Ala Gly Arg Leu Lys Pro Asp Glu
 385 390 395 400
 Gly Gly Glu Val Pro Val Leu Asn Val Ser Tyr Lys Pro Gln Lys Ile
 405 410 415
 Ser Pro Lys Ser Thr Gly Ser Val Arg Gln Leu Leu His Glu Lys Ile
 420 425 430
 Arg Asp Ala Tyr Thr His Pro Gln Phe Val Thr Asp Val Met Lys Pro
 435 440 445
 Leu Gln Ile Glu Asn Ile Ile Asp Gln Glu Val Gln Thr Leu Ser Gly
 450 455 460
 Gly Glu Leu Gln Arg Val Arg Leu Arg Leu Cys Leu Gly Lys Pro Ala
 465 470 475 480
 Asp Val Tyr Leu Ile Asp Glu Pro Ser Ala Tyr Leu Asp Ser Glu Gln
 485 490 495
 Arg Leu Met Ala Ala Arg Val Val Lys Arg Phe Ile Leu His Ala Lys
 500 505 510
 Lys Thr Ala Phe Val Val Glu His Asp Phe Ile Met Ala Thr Tyr Leu
 515 520 525
 Ala Asp Arg Val Ile Val Phe Asp Gly Val Pro Ser Lys Asn Thr Val
 530 535 540
 Ala Asn Ser Pro Gln Thr Leu Leu Ala Gly Met Asn Lys Phe Leu Ser
 545 550 555 560
 Gln Leu Glu Ile Thr Phe Arg Arg Asp Pro Asn Asn Tyr Arg Pro Arg
 565 570 575
 Ile Asn Lys Leu Asn Ser Ile Lys Asp Val Glu Gln Lys Lys Ser Gly
 580 585 590
 Asn Tyr Phe Phe Leu Asp Asp